環境を動的に変化させる戦略を持つ主体を構成要素として持つゲーム世界における戦略主体の進化と環境変遷ダイナミクス(生命・進化・ゲーム,基研長期研究会「複雑系4」)

秋山, 英三; 金子, 邦彦

研究会報告

環境を動的に変化させる戦略を持つ主体を構成要素として持つゲーム世界における戦略主体の進化と環境変遷ダイナミクス

東大教養　　秋山　英三[1]　　金子　邦彦[*]

1. はじめに

1.1. 概要

ゲーム環境自体がプレーヤーとプレーヤーの採る戦略と影響しあいながら発生・進化していく系について、その進化のダイナミクスと相互作用について、計算機のシミュレーションによる発見的手法を通して考察することが本研究の主な目的である。

生物進化、社会進化など主体間の戦略がお互いに影響を与えあう系において、進化の過程一般をゲームとして捉え、進化系をコンピュータープログラムのモデルに還元することによって進化の過程一般の理解を目指す、といった試みは過去からも多く行われてきた[1]。しかし、そこで設定として用いられるゲームは（従来のフォンノイマン以来の理論数学としてのゲーム理論でもそうだが）固定的なものが主に使われる[1]。つまり、各戦略主体の行動や価値体系の変化によって変動することがあってしかるべき利得関数が変動しない（Static Game）。また、ここで考察の中心となるのは主にプレーヤー間の相互作用（協力・裏切り・結託・共進化など）、またはプレーヤーの戦略そのものであった。

しかし、現実世界のゲーム環境は決して固定されてはいない。また、戦略主体と「ゲーム環境」自体との間の相互作用も大きな問題となりうる。例えば現実の生態系型システムでは、ある個体が採用する戦略が外部環境に変化を引き起こしてゲームの利得行列自体が変わってしまったり、その結果生じた新しい環境に、逆に個体が適応することが要請されたりする場合もある。また例えば、ある2個体間のローカルな利得行列が無数に存在する第3者達が取る戦略によって変動する場合もあるし、自分自身の内的変化によって同じ戦略に対する効用が異なってくる場合もある。（同じ行動に飽きてしまって喜びが少なくなる場合など。）そもそも、まったくゲーム（競争関係）が無かった所に外的要因からゲームが発生して来る場合もある。後述のモデルでも以上のような問題を取り扱っていく。

1.2. Dynamical Game Model

以上のように現実には普く存在しながら従来ほとんど取り扱われることが無かった系、つまり、ゲーム環境自体がプレーヤーの内的状態・プレーヤーの採る戦略と影響しあいながら発生・進化していく

---

[1] E-mail : akiyama@cyber.c.u-tokyo.ac.jp

[*] E-mail : kaneko@cyber.c.u-tokyo.ac.jp

1逐次ゲーム（破産ゲーム）、微分ゲーム（戦争ゲーム）などゲームが変化するモデルもある。ただし、その目的は主に方程式を解くことによって得られる最適均衡戦略を求めることである

系についての問題を取り扱うためのツールとして、新しいゲーム(Dynamical Game)のモデル（後述）を提唱したい。現実の系の現象の中でこのモデルが理論的基礎となりうるものはある程度限定されるが、少なくとも動的に変動する環境と戦略主体との関わりを考える上で、一つのきっかけとなり得るだろう。

1.3. 今回の内容

現時点では戦略のインプリメントの方法が決まってないので、シミュレーションは行われていない。今回は主にモデルの説明と今後着目していく点について見ていく。

2. モデルの構成

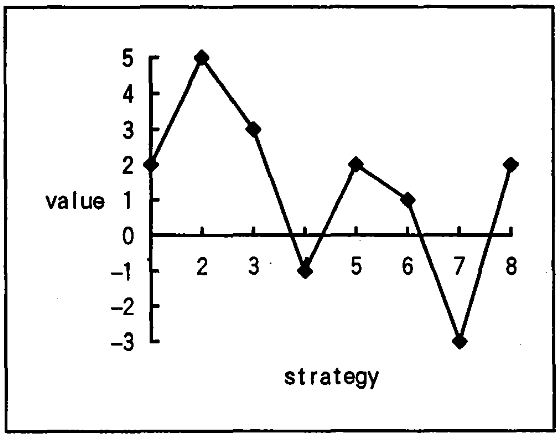
2.1. ゲームの環境の表現法

2.1.1. ゲームの利得行列とゲーム環境(Static Game)

Dynamical Game のモデルの紹介に入る前に、従来のゲーム理論のモデルについて簡単に復習しておく。下の表は一人ゲームの利得行列の一例である。

一人非ゼロ和ゲームの例

	戦略	環境 1
Player	*1*	+2
	2	+5
	3	+3
	4	-1
	5	+2
	6	+1
	7	-3
	8	+2



戦略１を選択すればプレーヤーは得点＋２を、戦略４を選択すれば、プレーヤーは得点－１を獲得できる。環境側は戦略１しか選択できない。もし、この行列が固定したまま変化しないなら、プレーヤーは戦略２を選ぶことは自明であり、古典ゲーム論的な立場からみれば単純なゲーム。最大化・最小化問題の範疇として取り扱われる（つまり、利得行列の発見する課程を問題にしている）。ここで、この利得行列で表現されている得点には好むと好まざるに関わらず、必ずプレーヤーの視点が入らざるを得ないことに留意しておきたい。つまり、客観的環境が同じでもプレーヤーにとっての環境はそのときそのときのプレーヤー自身の状態、価値観によって違ってくる。利得行列には環境自体の構造の他に、プレーヤーの内部状態が反映される。プレーヤーにとっての環境はプレーヤーの内部状態と切り離せない。

2.1.2. 2人非ゼロ和ゲームにおける環境(ゲーム利得行列)

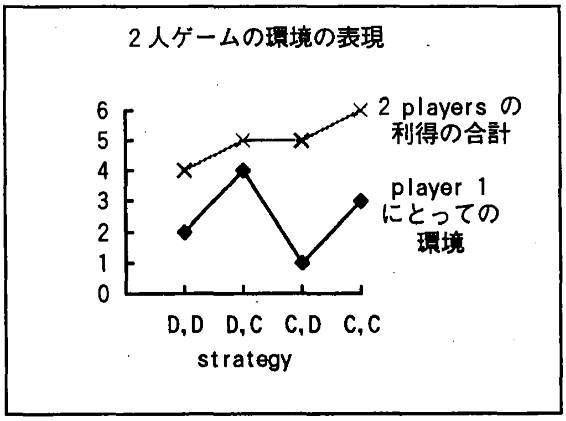
囚人ジレンマの利得行列

		player 2: *C*	player 2: *D*
player 1	*Co-operate*	3, 3	1, 4
	Defect	4, 1	2, 2

上は有名にな囚人ジレンマの(2-person non-zero-sum Game)利得行列だが、環境に注意を向けた表現をすると下のようになる。このような拡張によって、一人ゲームのモデルはN人ゲームのモデルに拡張することができる。

囚人ジレンマの利得行列(環境に着目した表現)

player 1, 2	環境(for player 1)	2 players の合計
D, D	+2	+4
D, C	+4	+5
C, D	+1	+5
C, C	+3	+6



2.2. 動的「一人」ゲームのモデル

von Neumann のゲーム理論と同じく、動的ゲームについても、極力シンプルなモデルを構成することが広い応用へとつながるものと思われる。以下、今後の研究で用いる Dynamical Game のモデルの紹介を行う。やはり本来は、N人ゲームのモデルに面白い現象が見られるはずであるが、とりあえずN人ゲームへの拡張は今後の課題ということにして、最初は一番シンプルな一人ゲームのモデルにおける現象から調べていくことにする。モデルの例としては様々なものがあるが、ここでは説明としてわかりやすい「差分モデル」について紹介する（2.2.3.節）。

2.2.1. 純粋戦略モデル

	H(t)	環境: *S1*
player	*1*	$u_1(t)$
	2	$u_2(t)$
	3	$u_3(t)$
	4	$u_4(t)$

◆ $\vec{u}(t+1) = f\left(\vec{u}(t),\ H(t)\right)$

環境もプレーヤーの選択した戦術に応じて変化する。この関数 $f(u,H)$ はそのゲーム世界での物理法則である。(こう動いたら環境にこういう影響がでる。)

◆ $E(t) = u_i(t) \cdots$ if $H(t) = i$

H(t)	時点 t でのプレーヤーの手の選択
u(t)	縦ベクトルで、その第 i 要素 u(i,t)は、時点 t で戦術 i を一単位することにより時点 t でプレーヤーが獲得するはずの効用を表す。
E(t)	時点 t にプレーヤーが獲得する効用

2.2.2. 混合戦略モデル

♦ $\vec{u}(t+1) = f\left(\vec{u}(t),\ H(t)\right)$

♦ $E(t) = \vec{x}(t) \cdot \vec{u}(t)$

x(t)	縦ベクトル。その第 i 要素 x(i,t)は、時点 t でなされる行動 i の単位数を表す。
u(t)	縦ベクトルで、その第 i 要素 u(i,t)は、時点 t で行動 i を一単位することにより時点 t でプレーヤーが獲得する効用を表す。
E(t)	u(t)と x(t)の内積として定義されるスカラーで、時点 t で得られる効用を表す。

2.2.3. 動的一人ゲームの簡単な例・・差分モデル

環境

	H(t)	
player	*1*	$u_1(t)$
	2	$u_2(t)$
	3	$u_3(t)$
	4	$u_4(t)$

♦ $E(t) = u_i(t) \quad \cdots$ if $H(t) = i$

♦ $\begin{cases} \vec{u}(t+1) = \vec{u}(t) + \Delta\vec{u} \\ \Delta\vec{u} = \text{行列}A\text{ の }H(t)\text{列成分} \end{cases}$

♦ $A = \begin{bmatrix} -2 & +1 & +3 & +2 \\ +2 & -1 & +1 & +2 \\ +3 & +4 & -2 & +1 \\ +2 & +1 & +1 & -3 \end{bmatrix}$ $H(t) = 4$ なら $\Delta\vec{\mathrm{u}} = \begin{pmatrix} +2 \\ +2 \\ +1 \\ -3 \end{pmatrix}$

H(t)	時点 t でのプレーヤーの手の選択
u(t)	縦ベクトルで、その第 i 要素 u(i,t)は、時点 t で

	戦術iを一単位することにより時点tでプレーヤーが獲得するはずの効用を表す。
A	(m,m)行列で、その(i,j)成分は、時点 t で戦術 j を一単位することによる u(i,t)の増分を表す。
E(t)	時点tにプレーヤーが獲得する効用

3. モデルの特徴と今後の着目点

3.1. ゲーム環境自体を変動させる戦略

現実の系のように、ある主体が選択する戦略によって環境が変化する可能性がある場合、その主体にとって環境を自分に有利に変化させる戦略を選択することが本質的な問題になる。選択された戦略により環境自体が動的に変化していくことになるが、逆にこの時、主体のもつアルゴリズム自体も環境の変化に対応して進化(学習)適応していく必要が生じる。例えば、人間が住居を造る、田畑を作る、古代にシアノバクテリアが酸素のある環境を作り出す、などの戦略によって環境が変化したり、あるいは人間のようにものを造るいう戦略によって環境が劇的に変化すると、その変化は戦略主体自身に大きな影響を与える。環境の変化に伴って過去に有効だった戦略も有効でなくなり（ある戦略の有効性はそのとき戦略主体がいる環境に大きく依存する）、新たな戦略の進化が要求される。2.2.3.節のモデルでいうと、プレーヤーはある戦略を選択することによって、その時点の利得行列に従って得点を獲得できるが、単に一番得点が高い戦略を選択するだけでは最終的に利得行列全体が低得点になって長期的な視点から見て困ったことになる可能性もある。問題はいかに「良い」環境を築くことができる戦略選択の系列（パス）を見つけれるかである。ある程度良い環境を築くことができるような進化を行ったプレーヤーは、さらにその環境を保ちつつ、さらに良い環境を築くべく進化することが可能であろう。例えば文章を書くためにタイプライターを使う、という戦略を採ることで文章作成の効率を上げた人間は、そのタイプライター（キーボード）が存在する環境に適応しつつ、更に次はワープロを使うという戦略を採るようになる（キーボードを使いつつ、鉛筆のように文字の削除もできる）。ワープロがある環境に人間が適応進化してくると文章の書き方自体もまた変化して（思いついたことからどんどん書き込んで、文章を切り貼りして仕上げる）文章を更に効率よく編集できるようになる（逆に、その環境をキープしなければ、つまり、ワープロを使わなければ文章を書きづらくなる）。戦略の進化によって引き起こされた環境の変化は、結局その戦略を選択した主体自身に影響を与える。当初、予想も付かなかった方向の進化をしばしば促すことになる。

（注意）2.2.3.節のモデルは、一人ゲームの範囲では最適解はすぐに見つかってしまう。あくまでも説明の便のためであり、一人ゲームとしては差分系と違う関数系を用意する。（N人ゲームとしては面白いゲームとなる可能性がある。）

3.2. プレーヤーの内部状態の変化と環境の変動

同じ戦略でもプレーヤーの価値体系、内部状態の変化に伴って、その価値(利得、効用)が変わる。例えば単純な例では、同じボール遊びを行っていてもだんだんつまらなくなる、などがこれに当たる。あ

るプレーヤーにとっての利得行列は決して客観的環境のみによって決まるわけではなくプレーヤーの内部状態と切り離して考えることはできない。モデルとしては、環境を表す変数とプレーヤーの内部状態変数をリンクさせた概念的なモデルがRossler[2]によって提唱されている。前述のモデルでは現段階ではこの点を考慮していないが将来的には導入する予定である。

3.3. 環境の変動と他のプレーヤーへの影響（N人ゲームへの拡張）

既にモデルのところで述べたように、この動的ゲームでは、最終的にはN人ゲームモデルへの拡張を考えている。動的一人ゲームのモデルでは、ある戦略を選び続けて他の戦略に対応する利得を増加させて後でその戦略を選んで高い効用を得る、といったことが可能である。（食糧を食べるのを我慢する戦略を続けて腹を空かせて食糧の効用を増加させる・畑を休ませて土地を肥沃にしてから作物を植えるetc.）しかし、N人ゲームでは我慢して効用が大きくなるのを待っていたら、効用が適度に大きくなったところで第三者に奪われてしまう可能性がある。また、あるプレーヤーの行動によって他のプレーヤー達の間のローカルな利得行列が変化するため、たとえ直接対峙していなくても、環境を操作することで結局他のプレーヤーの行動を制御できる可能性がある。こういった性質を持つ動的N人ゲームのモデルは、細胞分化モデル[3]など、様々な現象の解析への応用が可能であると思われる。

---

[1] K.Lindgren. Artificial Life II, pages 295-312, 1991

[2] O.E.Rossler, "Fraiberg-Lenneberg Speech", Chaos, Solitons & Fractals Vol.1, No.1, pp.125-131, 1994

[3] K.Kaneko and T. Yomo, "Cell Division , Differentiation, and Dynamic Clustering", Physica 75 D, pp. 89-102, 1994